GNSS/Acoustics seafloor positioning simulations using a multi-observations least square inversion: Supplementary Material

Pierre Sakic[1,2,3,*], Valérie Ballu[1], and Jean-Yves Royer[3]

[1]CNRS & University of La Rochelle, Littoral Environnement et Sociétés, 2 Rue Olympe de Gouges, 17000 La Rochelle, France

[2]GFZ German Research Centre for Geosciences, Helmholtz-Zentrum Potsdam, Telegrafenberg, 14473 Potsdam, Germany

[3]CNRS & University of Brest, Laboratoire Géosciences Océan, Institut Universitaire Européen de la Mer, Rue Dumont d'Urville, 29280 Plouzané, France

[*]Corresponding author : pierre.sakic@gfz-potsdam.de

January 31, 2020

1 General Formulation of a Least Square Inversion

We have a n number of observations l, and we want to determine p unknown physical quantities x. We have $n > p$. We also have a function F with several variables.

$$F : \begin{array}{ccc} \mathbb{R}^p & \longrightarrow & \mathbb{R}^n \\ (x_1, \dots, x_p) & \longmapsto & F(x_1, \dots, x_p) = (l_1, \dots, l_n) \end{array} \tag{1}$$

The exact values $(\dot{x}_1, \dot{x}_2, \dots, \dot{x}_p)$ of $(x_1, x_2, \dots, x_p)$ are inherently inaccessible because the observations are tainted with errors. Thus, we try to estimate values close to $(\hat{x}_1, \hat{x}_2, \dots, \hat{x}_p)$. We introduce the notion of residuals v, which is the difference between the actual observations l_i (called *stochastic model*) and the theoretical values λ_i obtained by the model based on the estimated parameters $\hat{x}_i$ (called *functional model*) (Sillard, 2001). So that:

We have :

$$\forall i \in [\![1, n]\!] \quad v_i = l_i - f\left(\hat{x}_1, \dots, \hat{x}_p\right) = l_i - \lambda_i \tag{2}$$

Then, we impose a condition on the residuals : we want the quadratic sum of the latter $\sum_i^n v^2$ to be minimal. It is called the **least squares condition** (Legendre, 1805; Gauss, 1809).

We call $\mathbf{L}$ the vector of the actual observations (length n):

$$\mathbf{L} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{bmatrix} \tag{3}$$

We call $\mathbf{X}$ the vector of the unknown parameters to estimate (p length):

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} \tag{4}$$

And we call $\mathbf{V}$ is the vector of the residuals associated with each observation at the end of the inversion (length n):

$$\mathbf{V} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \tag{5}$$

The problem must be linearized near a solution close enough to $(\dot{x}_1, \dot{x}_2, \dots, \dot{x}_p)$: this solution is called the *solution a priori* and write it $\mathbf{X_0} = (x_{0,1}, x_{0,2}, \dots, x_{0,p})$. To do this, we introduce the matrix of *partial derivatives* or *Jacobian* $\mathbf{J}$ of the function F, of size (n, p) :

$$\mathbf{J_F}(\mathbf{X}) = \begin{bmatrix} \dfrac{\mathrm{d}f_1}{\mathrm{d}x_1} & \cdots & \dfrac{\mathrm{d}f_1}{\mathrm{d}x_p} \\ \vdots & \ddots & \vdots \\ \dfrac{\mathrm{d}f_n}{\mathrm{d}x_1} & \cdots & \dfrac{\mathrm{d}f_n}{\mathrm{d}x_p} \end{bmatrix} \tag{6}$$

And we call $\mathbf{A}$ the matrix of partial derivatives *in the neighborhood* of $\mathbf{X_0}$, also called *design matrix.*

We have $\mathbf{A} = \mathbf{J_F}(\mathbf{X_0})$

We call $\mathbf{\Lambda}$ the vector of modeled observations associated with *a priori* values such as:

$$F(\mathbf{X_0}) = \mathbf{\Lambda} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} \tag{7}$$

We set $\mathbf{B}$ the vector of the differences between the actual observations l_i and the modeled observations associated with the *a priori* values λ_i. We have $\mathbf{B} = \mathbf{L} - \mathbf{\Lambda}$.

Lastly, we introduce the notion of weight, which indicates the quality of the observations, and allows to homogenize them if they are of different nature: each observation b_i is associated to a standard deviation ς_i quantifying its accuracy. The more accurate the measurement, the lower the standard deviation. We then define for each ς_i weight π_i such that $\pi_i = \dfrac{1}{\varsigma_i^2}$. The more reliable the measurement, the greater the weight. The associated matrix called *weights* is a diagonal matrix [1], such as :

$$\mathbf{P} = \begin{bmatrix} \frac{1}{\varsigma_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\varsigma_2^2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \frac{1}{\varsigma_n^2} \end{bmatrix} = \begin{bmatrix} \pi_1 & 0 & \dots & 0 \\ 0 & \pi_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \pi_n \end{bmatrix} = \begin{bmatrix} \frac{1}{\varsigma_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\varsigma_2^2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \frac{1}{\varsigma_n^2} \end{bmatrix} \tag{8}$$

If the weight matrix is considered, the least squares condition becomes:

$$\sum_i^n \frac{v_i^2}{\sigma_i^2} = \mathbf{V^tPV} \text{ minimal} \tag{9}$$

And the optimal solution in the least squares sense $\hat{\mathbf{X}} = \mathbf{X_0} + \boldsymbol{\delta}\mathbf{X}$, where $\boldsymbol{\delta}\mathbf{X}$ is the correction to make to the *a priori* initial estimate, is estimated by the following system:

$$\begin{cases} \mathbf{A}\boldsymbol{\delta}\mathbf{X} + \mathbf{V} = \mathbf{B} \\ \mathbf{V^tPV} \text{ minimal} \end{cases} \tag{10}$$

If we call the matrix $\mathbf{N} = \mathbf{A^tPA}$ the *normal matrix*, the correction $\boldsymbol{\delta}\mathbf{X}$ is obtained by solving the *normal equation*:

$$\boldsymbol{N}\boldsymbol{\delta}\mathbf{X} = \mathbf{A^tPB} \tag{11}$$

Which gives :

$$\boxed{\boldsymbol{\delta}\mathbf{X} = \mathbf{N^{-1}A^tPB}} \tag{12}$$

The residues are obtained as follows:

$$\hat{\mathbf{V}} = \mathbf{B} - \mathbf{A}\hat{\mathbf{X}} = \mathbf{L} - f(\hat{\mathbf{X}}) \tag{13}$$

Following an iterative process, the new estimate $\hat{\mathbf{X}}$ becomes the new *a priori* $\mathbf{X_0}$ in the next inversion step. We have, at the k -th iteration of the inversion:

$$\hat{\mathbf{X}}_k = \mathbf{X}_{\mathbf{0},k+1} \tag{14}$$

We stop the iterations when the convergence criterion is fulfilled, *i.e.* when $\|\boldsymbol{\delta}\mathbf{X}_k\| < \kappa$ (where κ is a predetermined threshold).

1.1 About the observation function

The function F is an *ad hoc* multivariate function of $\mathbb{R}^p \longrightarrow \mathbb{R}^n$, specific to each problem. It can be separated into n *observation equations* f_i.

For each observation l_i, we have the associated functional model:

$$\lambda_i = f_i(x_1, \dots, x_p, \boldsymbol{\Omega}_{l_i}) \tag{15}$$

where $\boldsymbol{\Omega}$ is a set of additional observations.

[1]in the simplified case where the uncorrelated observations are assumed

We can then define F:

$$F(x_1,\dots,x_p) = \begin{cases} f_1(x_1,\dots,x_p,\boldsymbol{\Omega}_{l_1}) & = & \lambda_1 \\ \vdots & & \\ f_n(x_1,\dots,x_p,\boldsymbol{\Omega}_{l_n}) & = & \lambda_n \end{cases} \tag{16}$$

We call the function F, common to all observations, the *observation function.*

1.2 Constraints from the Helmert Method

In our case, some additional information links the unknowns to one another. Thus, it is necessary to augment the system with additional equations. We use the constraining method described by the German geodesist Friedrich Helmert (Helmert, 1872). It allows to set some unknowns with predefined values, or to specify some relations linking the unknowns to one another.

It aims to "surround" the normal matrix $\mathbf{N}$ with a matrix $\mathbf{C}$ representing the constraints on the parameters (Ghilani, 2011):

$$\begin{bmatrix} \mathbf{N} & \mathbf{C}^{\mathrm{t}} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta}\mathbf{X} \\ \boldsymbol{\Gamma} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{\mathrm{t}}\mathbf{PB} \\ \boldsymbol{\Phi} \end{bmatrix} \tag{17}$$

If there are q constraint equations, $\mathbf{C}$ is a $q \times p$-sized matrix that describes the relationships between p parameters, and $\boldsymbol{\Phi}$ is a vector of length q where the constraints values are stored. $\boldsymbol{\Gamma}$ denotes the vector of Lagrange multipliers, estimated in addition to $\boldsymbol{\delta}\mathbf{X}$.

This new normal equation is solved in the same way as the classical version, by inverting the augmented normal matrix.

2 Generic design matrix definition

If n_R transponders are used, the associated design matrix $\mathbf{A}_{SMA}$ is block-diagonal, of size $(n_\tau \cdot n_R, 3n_R)$, in the ideal case where each transponder has responded to n_τ pings sent from the surface. In a more realistic case, its size is $(\sum_{i_R=1}^{n_R} n_{\tau,R_i}, 3n_R)$ where we have a number n_{τ,R_i} of pings for each receiver R_i

$$\mathbf{A}_{SMA} = \begin{bmatrix} \mathbf{A}_{SMA,R_1} & 0 & \cdots & 0 \\ 0 & \mathbf{A}_{SMA,R_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{A}_{SMA,R_{n_R}} \end{bmatrix} \tag{18}$$

with :

$$\mathbf{A}_{SMA,R_i} = \begin{bmatrix} \frac{\mathrm{d}f_{SMA,\tau_1}}{\mathrm{d}x_{R_i}} & \frac{\mathrm{d}f_{SMA,\tau_1}}{\mathrm{d}y_{R_i}} & \frac{\mathrm{d}f_{SMA,\tau_1}}{\mathrm{d}z_{R_i}} \\ \frac{\mathrm{d}f_{SMA,\tau_2}}{\mathrm{d}x_{R_i}} & \frac{\mathrm{d}f_{SMA,\tau_2}}{\mathrm{d}y_{R_i}} & \frac{\mathrm{d}f_{SMA,\tau_2}}{\mathrm{d}z_{R_i}} \\ \vdots & \vdots & \vdots \\ \frac{\mathrm{d}f_{SMA,\tau_{n_\tau}}}{\mathrm{d}x_{R_i}} & \frac{\mathrm{d}f_{SMA,\tau_{n_\tau}}}{\mathrm{d}y_{R_i}} & \frac{\mathrm{d}f_{SMA,\tau_{n_\tau}}}{\mathrm{d}z_{R_i}} \end{bmatrix} \tag{19}$$

f_{SMA,τ_i} is the observation function associated to a two-way travel time τ_i as defined in relation 2 of the main article. It can be differentiated numerically using a method described in Abramowitz and Stegun

(1965) or Burden and Faires (1998) for instance. In our study, we implemented the *three-point midpoint formula.*

To each acoustic observation, a weight $\pi_{\tau,i}$ is associated, to complete the weight matrix $\mathbf{P}_{SMA}$.

3 Baseline length observations

Starting from the observation function 15 of the main article, we have the following derivatives :

$$\frac{\mathrm{d}f_D(\mathbf{X}_{R_A},\mathbf{X}_{R_B})}{\mathrm{d}\mathbf{X}_{R_A}} = \left[\frac{x_{R_A}-x_{R_B}}{D_{AB}}, \frac{y_{R_A}-y_{R_B}}{D_{AB}}, \frac{z_{R_A}-z_{R_B}}{D_{AB}}\right] \tag{20}$$

et

$$\frac{\mathrm{d}f_D(\mathbf{X}_{R_A},\mathbf{X}_{R_B})}{\mathrm{d}\mathbf{X}_{R_B}} = \left[\frac{x_{R_B}-x_{R_A}}{D_{AB}}, \frac{y_{R_B}-y_{R_A}}{D_{AB}}, \frac{z_{R_B}-z_{R_A}}{D_{AB}}\right] \tag{21}$$

The observation vector $\mathbf{L}$ is enhanced so as :

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_{SMA} \\ \mathbf{L}_D \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \vdots \\ \tau_{n_\tau} \\ D_{12} \\ \vdots \\ D_{ij} \\ \vdots \\ D_{n_R-1,n_R} \end{bmatrix} \tag{22}$$

And the design matrix is also enhanced :

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{SMA} \\ \mathbf{A}_D \end{bmatrix} \tag{23}$$

where $\mathbf{A}_{SMA}$ is the design matrix defined in the main article section 2.3, and $\mathbf{A}_D$ the design matrix of baseline length measurements, of size $(n_D, 3n_R)$.

We have (with $n = n_R$):

$$_D = \begin{bmatrix} \frac{x_{R_1}-x_{R_2}}{D_{12}} & \frac{y_{R_1}-y_{R_2}}{D_{12}} & \frac{z_{R_1}-z_{R_2}}{D_{12}} & \frac{x_{R_2}-x_{R_1}}{D_{21}} & \frac{y_{R_2}-y_{R_1}}{D_{21}} & \frac{z_{R_2}-z_{R_1}}{D_{21}} & 0 & \cdots 0 \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 \cdots & \frac{x_{R_i}-x_{R_j}}{D_{ij}} & \frac{y_{R_i}-y_{R_j}}{D_{ij}} & \frac{z_{R_i}-z_{R_j}}{D_{ij}} & \cdots 0 \cdots & \frac{x_{R_j}-x_{R_i}}{D_{ji}} & \frac{x_{R_j}-x_{R_i}}{D_{ji}} & \frac{x_{R_j}-x_{R_i}}{D_{ji}} & \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots 0 \cdots & 0 & \frac{x_{R_{n-1}}-x_{R_n}}{D_{n-1,n}} & \frac{y_{R_{n-1}}-y_{R_n}}{D_{n-1,n}} & \frac{z_{R_{n-1}}-z_{R_n}}{D_{n-1,n}} & \frac{x_{R_n}-x_{R_{n-1}}}{D_{n,n-1}} & \frac{x_{R_n}-x_{R_{n-1}}}{D_{n,n-1}} & \frac{x_{R_n}-x_{R_{n-1}}}{D_{n,n-1}} \end{bmatrix} \tag{24}$$

Or, in a simplified form :

$$\mathbf{A}_D = \begin{bmatrix} \mathbf{A}_{D_{12}} & \mathbf{A}_{D_{21}} & 0 & \cdots 0 \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 \cdots & \mathbf{A}_{D_{ij}} & \cdots 0 \cdots & \mathbf{A}_{D_{ji}} & \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots 0 \cdots & 0 & \mathbf{A}_{D_{n-1,n}} & \mathbf{A}_{D_{n,n-1}} \end{bmatrix} \tag{25}$$

With :

$$\mathbf{A}_{D_{ij}} = \begin{bmatrix} \frac{x_{R_i}-x_{R_j}}{D_{ij}} & \frac{y_{R_i}-y_{R_j}}{D_{ij}} & \frac{z_{R_i}-z_{R_j}}{D_{ij}} \end{bmatrix} \tag{26}$$

It is also necessary to add weights $\pi_{D,i}$ corresponding to the lengths of baselines in the corresponding matrix P, since we add observations of different nature to the problem.

4 Direct estimation of the barycentrer coordinates

Relation 9 in the main article stipulates that :

$$\sum_{i=1}^{n_R} \Delta\mathbf{X}_{R_i} = 0 \tag{27}$$

meaning that the sum of the coordinate differences between all transponders must be equal to zero.

For n transponders on the seafloor, the vector of unknowns is :

$$\mathbf{X} = \begin{bmatrix} x_G \\ y_G \\ y_G \\ \Delta x_{R_1} \\ \Delta y_{R_1} \\ \Delta z_{R_1} \\ \vdots \\ \Delta x_{R_n} \\ \Delta y_{R_n} \\ \Delta z_{R_n} \end{bmatrix} \tag{28}$$

Since

$$\frac{\mathrm{d}f_{SMA,\tau}}{\mathrm{d}\mathbf{X}_G} = \frac{\mathrm{d}f_{SMA,\tau}}{\mathrm{d}\Delta\mathbf{X}_R} \tag{29}$$

by analogy with the relations 18 and 19, the design matrix $\mathbf{A}_{SMA}$ takes the following form:

$$\mathbf{A}_{SMA} = \begin{bmatrix} \mathbf{A}_{SMA,R_1} & \mathbf{A}_{SMA,R_1} & 0 & \cdots & 0 \\ \mathbf{A}_{SMA,R_2} & 0 & \mathbf{A}_{SMA,R_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{SMA,R_{n_R}} & 0 & 0 & \cdots & \mathbf{A}_{SMA,R_{n_R}} \end{bmatrix} \tag{30}$$

Following the Helmert's method described in section 1.2 (equation 17), the matrix $\mathbf{C}$ and the vector $\mathbf{\Phi}$ are respectivively equal to :

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & \cdots & \underbrace{0 & 0 & 1}_{n_R \text{times}} & \cdots \end{bmatrix} \tag{31}$$

$$\mathbf{\Phi} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{32}$$

5 Estimation of a single depth

If we assume a single common depth for all seafloor transponders, then the vector of unknowns becomes :

$$\mathbf{X} = \begin{bmatrix} x_{R_1} \\ y_{R_1} \\ \vdots \\ x_{R_n} \\ y_{R_n} \\ \bar{z} \end{bmatrix} \tag{33}$$

The design matrix takes the following shape :

$$\mathbf{A}_{SMA} = \begin{bmatrix} \mathbf{A}_{SMA,R_1} & 0 & \cdots & 0 & \frac{\mathrm{d}f_{SMA,\tau_1}}{\mathrm{d}\bar{z}} \\ 0 & \mathbf{A}_{SMA,R_2} & \cdots & 0 & \frac{\mathrm{d}f_{SMA,\tau_2}}{\mathrm{d}\bar{z}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{A}_{SMA,R_{n_R}} & \frac{\mathrm{d}f_{SMA,\tau_{n_\tau},R_{n_R}}}{\mathrm{d}\bar{z}} \end{bmatrix} \tag{34}$$

In this case:

$$\mathbf{A}_{SMA,R_i} = \begin{bmatrix} \frac{\mathrm{d}f_{SMA,\tau_1}}{\mathrm{d}x_{R_i}} & \frac{\mathrm{d}f_{SMA,\tau_1}}{\mathrm{d}y_{R_i}} \\ \frac{\mathrm{d}f_{SMA,\tau_2}}{\mathrm{d}x_{R_i}} & \frac{\mathrm{d}f_{SMA,\tau_2}}{\mathrm{d}y_{R_i}} \\ \vdots & \vdots \\ \frac{\mathrm{d}f_{SMA,\tau_{n_\tau},R_i}}{\mathrm{d}x_{R_i}} & \frac{\mathrm{d}f_{SMA,\tau_{n_\tau},R_i}}{\mathrm{d}y_{R_i}} \end{bmatrix} \tag{35}$$

6 Depth differences as observables in the least-squares sense

If we arbitrarily consider the transponder R_1 and its depth z_{R_1} as a depth reference, we can enhance the observation vector $\mathbf{L}$ with depth differences. Thus, we have :

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_{SMA} \\ \mathbf{L}_z \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \vdots \\ \tau_{n_\tau} \\ \delta z_{12} \\ \vdots \\ \delta z_{1j} \\ \vdots \\ \delta z_{1,n_R} \end{bmatrix} \tag{36}$$

The design matrix, as described in equation 12 of the main article, is concatenated with the binary array $\mathbf{A}_z$ of size $(n_R, 3 + 3n_R)$:

$$\mathbf{A}_z = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & & 0 & 0 & 0 & & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & & & \ddots & \vdots & \vdots & \vdots & & \mathbf{0} & \\ 0 & 0 & 0 & 0 & 0 & -1 & & & & & 0 & 0 & 1 & & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \mathbf{0} & & & \vdots & \vdots & \vdots & \ddots & & \\ 0 & 0 & 0 & 0 & 0 & -1 & & & & & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{37}$$

The column of the component z_{R_1} of the transponder taken as depth reference (here from transponder R_1) is filled with coefficients -1. The elements corresponding to the observation $\delta z_{1,j}$ and the vertical component z_{R_j} of each transponder R_j are equal to 1. It is also necessary to introduce a specific π_z weighting the depth difference observations.

References

Abramowitz, M. and Stegun, I. A. (1965). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation.

Burden, R. and Faires, J. D. (1998). *Numerical Methods, Brooks.*

Gauss, C. F. (1809). *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium.* Göttingen.

Ghilani, C. (2011). Adjustment computations: spatial data analysis. *International Journal of Geographical Information Science*, 25(2):326–327.

Helmert, F. R. (1872). *Die Ausgleichungsrechnung nach der Methode der kleinsten Quadrate: mit Anwendungen auf die Geodäsie und die Theorie der Messinstrumente*, volume 1. Teubner, Leipzig.

Legendre, A. M. (1805). *Nouvelles méthodes pour la détermination des orbites des comètes.* Number 1. F. Didot, Paris.

Sillard, P. (2001). *Estimation par moindres carrés.* Hermès science publications.